EPIQ IQC, PCE-I-00-96-00002-00, Task Order No.839

# Atelier de zonage du corridor forestier Ranomafana-Andringitra-Ivohibe

## Lessons learned for the use and development of the program Plateforme d'Analyse for local priority areas analysis

*Prepared by:*

David Lees
Consultant, Projet d'Appui à la gestion de l'Environnement
International Resources Group, Ltd.
1211 Connecticut Avenue, NW, Suite 700
Washington, DC 20036 USA

*Prepared for:*
USAID/Madagascar

Date
January 2002

Environment Policy and Institutional Strengthening Indefinite Quantity Contact (EPIQ)

Partners: International Resources Group and Winrock International

Subcontractors: PADCO, Management Systems International, and Development Alternatives, Inc.

Collaborating Institutions: Center for Naval Analysis Corporation, Conservation International, KNB Engineering and Applied Sciences, Inc., Keller-Bliesner Engineering, Resource Management International, Inc., Tellus Institute, Urban Institute, and World Resources Institute

# Contents

# Introduction

A workshop was held from 19-21 December 2001 to refine a draft of the prioritisation of the Ranomafana-Andringitra-Ivohibe corridor at regional level by integrating biological and socio-economic data.

This report addresses the performance of Plateforme d'Analyse (PDA) in prioritising biodiversity data at such a regional level (here, for Fianarantsoa Province). PDA is in the first instance a prototype database and mapping program developed by Wildlife Conservation Society and Center for Conservation Biology at Stanford University that is well placed to serve as the cornerstone for the development of a national Malagasy program in biodiversity data synthesis and distribution via the world wide web. PDA is more than a mere computer program, since it initiates a process for biodiversity data sharing as conceived in the first workshop held at ONE in Antanananarivo in March 2000. It also provides a platform (at present implemented in prototype versions in Arcview and Visual Basic) for the simultaneous analysis of biodiversity data. The main uses of such analyses are likely to be for rapidly informing biodiversity planning and conservation evaluation exercises, environmental impact assessments, and querying threat status (e.g. establishing IUCN/CITES status). This is the first time the program has been seriously evaluated though at such a fine, local scale. Since a critical workshop in 1995 was held in Antananarivo that brought together key experts in biodiversity and socio-economic studies to identify the priorities for conservation and further research at a national level, local priority-setting workshops have started to come into vogue. It is important that PDA or its successor is up to speed with them.

# Preparing the available data

The data chosen and available for this workshop was birds, mammals and butterflies, almost all species in these groups (about 700 spp.), treated at a national level. The bird data was reformatted from a working Excel database (about 5000 records) supplied by ZICOMA; the mammals (about 7000 records) and butterflies (about 11000 records) were previously available from prior work for the PDA mainly compiled by this consultant, with additional input by consultants for PAGE and ZICOMA (see Fig. 1). The mammal and butterfly data had been previously cleaned, but it was decided that the mammal data should be substantially updated to reflect recent literature particularly for the Fianarantsoa province region. Due to time constraints it was necessary to reformat the ZICOMA data and add in the extra PDA-formatted mammal data for updated literature input during the workshop with the result that formatting and cleaning took up much of the available time and analytical work had to be done at night.

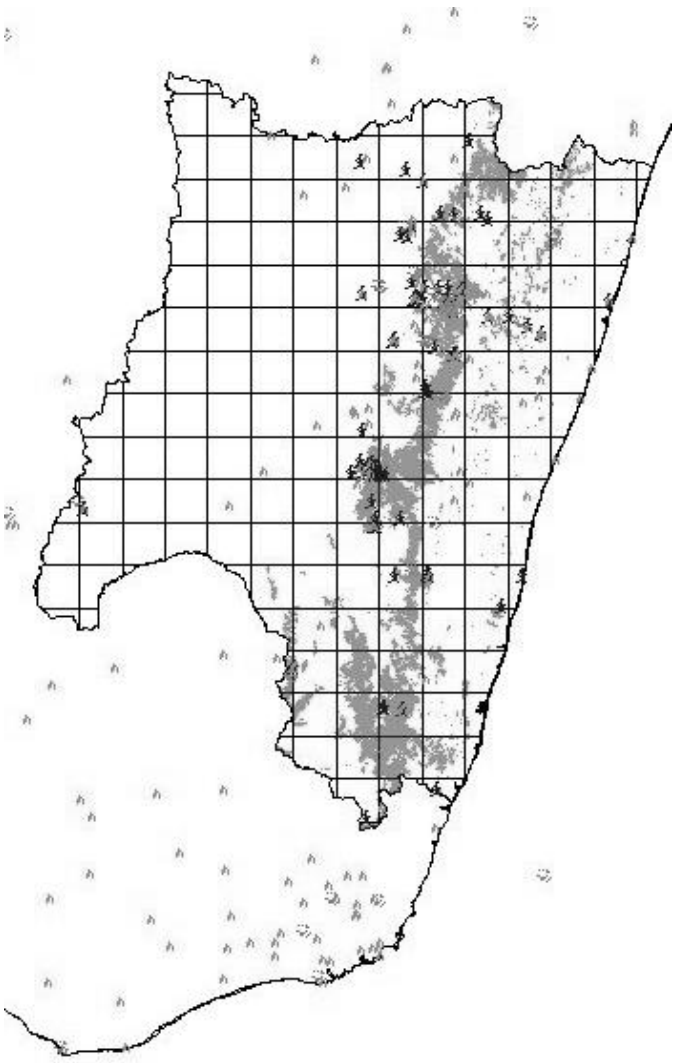

**Fig. 1.** Available data for mammals (blue), butterflies (pink) and birds (brown) overlain on the quarter-degree sided grid-cells (about 27 by 27 km, almost equal planar area) for Fianarantsoa Province. Green background: extent of humid forests in Fianarantsoa Province.

## Analysing the available data

The PDA extension in Arcview and Visual Basic currently has the following six principal functions:
Create Multiple Ranges
Browse Ranges
Create Diversity Surface
Create Endemic Clumps of Range
Query Parks for Intersecting Ranges (additional button)
Query Endemic Clumps for Intersecting ranges (additional button)

The Arcview version has the benefit of additional flexibility and functionality by dent of its integration within the inbuilt Arcview interface. It was this version we decided to use during the workshop.

For the Fianarantsoa exercise, it was decided to divide the province into quarter-degree sized gridcells (see Fig. 1).

It was decided not to use the plant data pre-formatted and provided by WWF for Andringitra, since no national context was available for the plant species. However, these data will be very useful in future as more botanical databases are incorporated. We used SQL to link the combined Excel file for mammals, birds and butterflies in PDA format (about 20Mb) with Arcview. This was successful and it was possible to view and interrogate the points on screen (see Fig. 1).

We then attempted to create multiple range files using PDA based first on range extents only (convex polygons), then based on clip with an integer elevation surface, and then using elevational range and categorical surface (in this case, habitat). Whilst generating the simple range extents is fairly fast, over about one hundred species across thousands of individual records cannot be processed in live time by PDA even using moderately fast laptops with adequate RAM (the ones at our disposal varied from 233-500 MHz and 160-256 Mb RAM). This is especially the case with more complex habitat shape files at 1:500,000 resolution or better such as the one supplied for the region for the workshop by MEF based on IEFN, or the BD500 shape file. PDA will definitely still work in these cases on cleaned data, but 1-5 minutes may need to be allowed for processing each species. All Madagascar mammals took something like 4-6 hours on the fastest laptop we had available. This is in principal OK though for overnight processing (with hard drives and screen set to always-on), except when an Arcview error dialog comes up, in which case user interaction is required. We also experienced computer freezes which were apparently due to excessive file sizes (20 Mb excel file increasing to 600 MB shapefile) and Windows dialogs came up requiring the virtual memory maximum file size to be resized on a laptop with 10Gb free!

There was not time to clean totally the mammal and bird data newly integrated in the workshop, which caused some errors. It proved infeasible to produce more than simple range extent files during the workshop using PDA, so some workarounds (intersecting the point database with the quarter-degree grid-cell references) were used to generate diversity data for each quarter degree grid-cell which was shown on the final morning of the workshop. Using the simple range extent files PDA is very fast (takes less than a few minutes) for producing expected maximum diversity surfaces at the required resolution, and so we could compare the species richness trend surface with the empirically recorded lists for each cell. Not all of these outputs can be generated at present using PDA itself.

The endemic clumps tool was illustrated during the workshop for birds but proved of limited use for estimation of areas of pre-defined endemism. A more flexible approach is to query the PDA-created range theme using the Arcview interface and then use PDA to generate a diversity surface on the endemic subset (or IUCN threatened species) subset. For example, range area instead can be used instead of range span. A result is illustrated below for some range restricted butterflies of the Fianarantsoa region (*Fulda imorina, Heteropsis cowani, H. fuliginosa, H.*

*narova, H. parvidens, Lepidochrysops azureus, Perrotia eximia, P. kingdoni, P. malchus, Strabena argyrina, S. daphne, S. sufferti*) all of which have national ranges less than or equal to a quarter of that possible (less than about 3.5 degrees latititude and longitude). A problem here is that very small or single point endemics (which may be the highest priority organisms for area

selection) often do not show up in the multiple polygon range list.

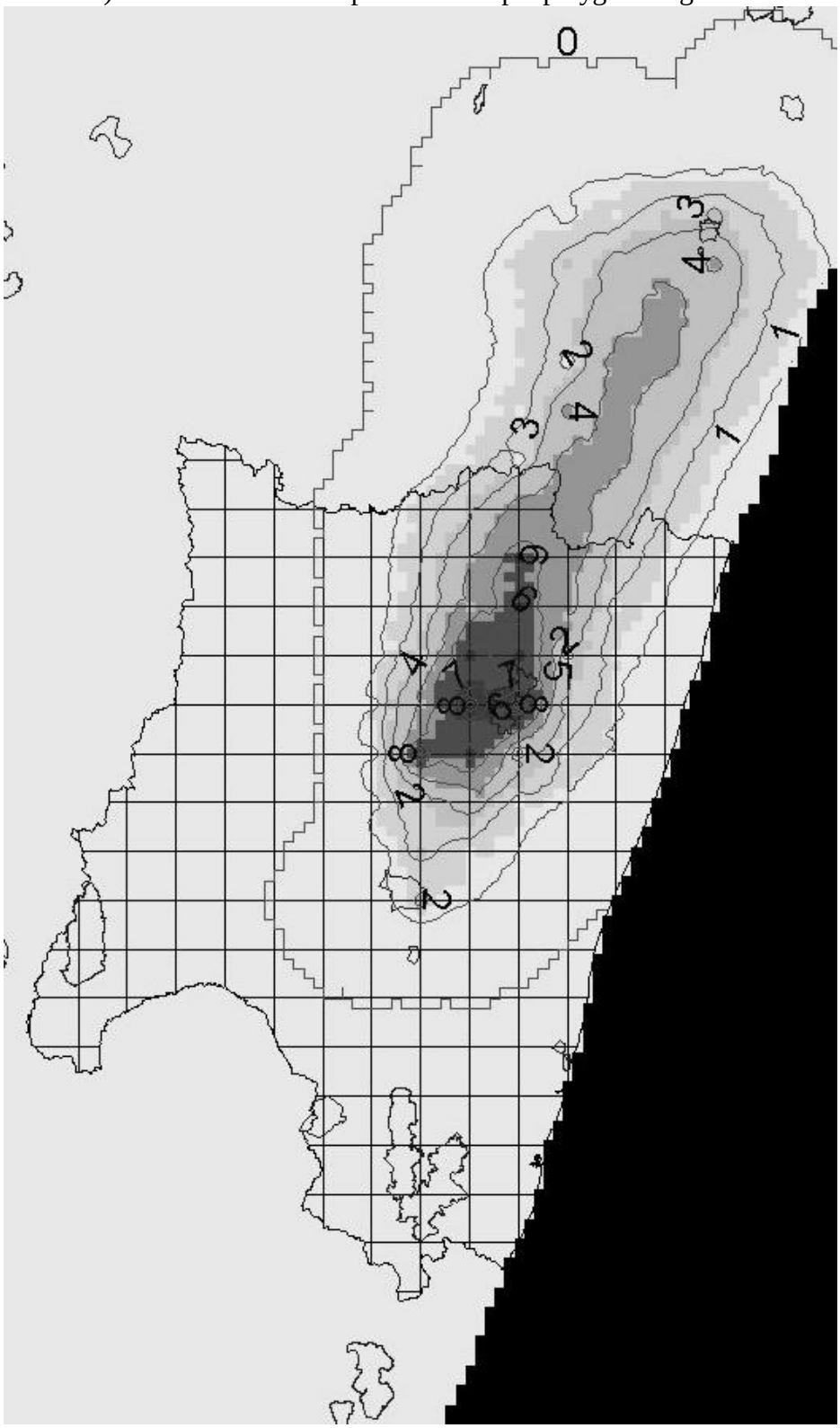


**Fig. 2.** Endemism surface for selected range restricted butterflies that occur in Fianarantsoa Province. The contours are number of endemic species. Note that Ranomafana National Park appears to protect an area of high endemism.

There was not time to use the parks query tool extensively. The query tool failed to interrogate the square polygons of the quarter degree grid even when we changed the name of the file to “Parks”. The clumps query tool produces anomalous results in such a case for which it was nmot designed anyway. In the next generation of PDA, it would be very useful to have a more generic polygon query tool above all with the facility for tabular output of results.

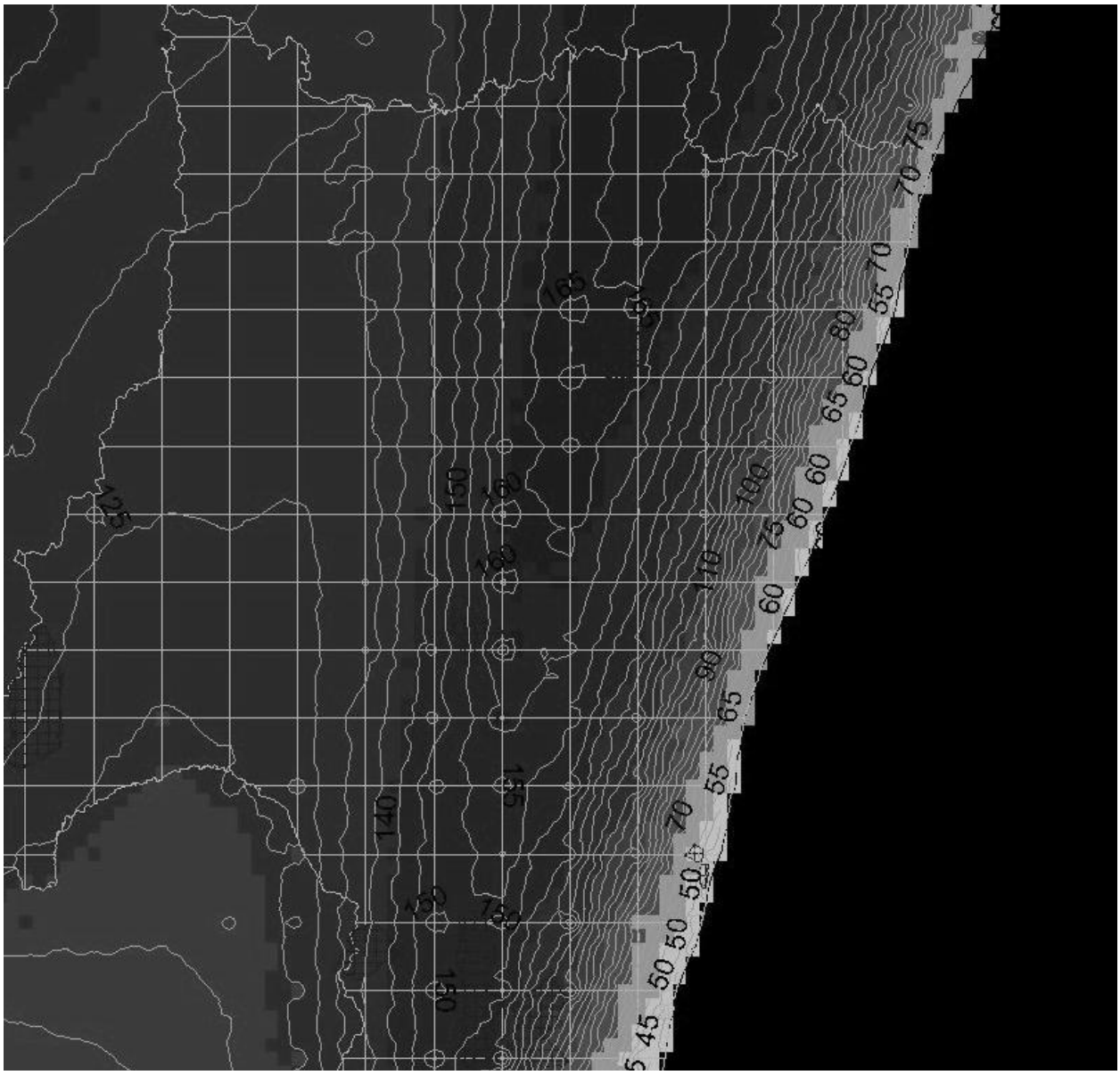


**Fig. 3.** Expected diversity gradient for birds and mammals across Fianarantsoa province based the overlay of range extent polygons in PDA.

## The most powerful current features of PDA

PDA has many powerful features that make it an invaluable asset for priority-setting exercises. I emphasise here just two of the most important ones:

1. Batch facility for generation of polygon range files from databases (best linked using SQL). The multiple range dialog is already very flexible since one can use a wide range of integer grid and categorical themes (not just habitat; one might want to use climate, soil, geology or many other themes within Arcview).
2. Diversity surface. As Figs 2 and 3 illustrate, the diversity surface interpolation within Arcview is not just fast but very flexible within the Arcview interface, for example if contours are then generated and labelled. One can even generate something approaching a probability surface if a diversity surface is drawn on a single species range.

## Summary of current limitations of PDA extension for priority areas analysis of biodiversity data

1. We found that the extension currently lacks full functionality in Laborde format when using the multiple range dialog. It is necessary to transfer all required files in advance to Geographic format. This can be done most easily using a suitably modified projection engine c:\esri\avgis_30\arcview\etc\default.prj or using the Prjctr! extension in the samples subdirectory with the appropriate conversion parameters (Projection: Hotine Oblique Mercator; Spheroid: International 1909; Longitude of central point: 46.43722917; Latitude of central point: -18.9; Azimuth of central line: 18.9; Scale factor along central line: 0.99950; False Easting: 1113136.3146; False Northing: 2882900.7279).

2. The extension does not presently query polygons that are not grids (lattices) nor endemic clumps for the taxa that occur or may occur within them. An obvious requirement for local priority-setting exercises is to be able to query each cell of a regularised grid such as the one shown in Fig. 1 for species with records within each cell, or with ranges that pass through each cell. It should be easy to modify the parks tool to accept polygons of any nature, including arbitrarily created ones. We tried this with the quarter degree polygon grid shapefile wmqd.shp in the workshop but without success.

3. The extension has no facility to output results from the park query. This type of export is essential for analysis.

4. A facility is needed to subset those ranges at a national level that pass through the local region selected. These are the only ranges that are needed for the local analyses. At present this can be done through an unnecessarily complicated procedure using the Geoprocessing Wizard in the Arcview interface (with Spatial Analyst loaded): (a) intersect the local shape using the geoprocessing wizard to produce a new shape file that shows the segments of the national ranges that pass through the [Fianar] province (Browse Range can be used in conjunction with the resulting file); (b) join the resulting file to the original (national) multiple polygon range file by opening both tables and performing a join by name; (c) start editing the joined table, sort by those species which have the local shape name appended and delete those records (species) that do not and (d) save as a new multiple range theme.

5. A facility is needed to select out those locally intersecting ranges that fall within certain criteria of range size or threat status/non-protected status. Those are the species that are most likely to discriminate areas for selection. The clumps tool does not exactly do this. At present such selections can be done on the result of the procedure in (4) using the query builder in Theme Properties and selecting a query like "<Width>" <= 3.5 and "<Height>" <= 3.5, then saving the query as a new theme.

6. A facility is needed to display rapidly the following types of results for a selected set of grid-cells

a) species richness (currently PDA will produce species diversity surfaces for overlain polygon ranges, but will not calculate the number of empirically recorded species nor the intersecting ranges within polygons such as a grid-cell lattice);
b) richness among species with particular endemic criteria;
c) turnover or complementarity among spatial neighbourhoods of gridcells that reflects the uniqueness of each grid-cell fauna or flora to other cells within its 9 or 25-cell neighbourhood);
d) gap analyses or generation of the deficit between the number of ranges intersecting grid-cells and the species actually recorded for grid-cells (this gives an idea of the likely species sampling deficit per grid-cell).

7. A facility is needed for joining sets of species range polygons together. Since processing time is a major limitation of PDA for large numbers of species, it is convenient to produce batches of species at a time and join these sets of polygons together later for a simultaneous diversity analysis.

8. A range is not produced for certain species, essentially those that only are known from one or two points. At present their presence can only be indicated in Browse Ranges when the corresponding database is also selected. However, it would be useful if even zero values were returned in the range theme table for these species so that they can be sorted/prioritised according to their status as the most endemic of endemic species.

9. Whilst extract values from the elevation theme works, we could not get automatic extraction of values from a habitat theme to work. In any case, for a species with many points, the virtual list of habitat types will rapidly become overly complex simply because of spatial and particularly temporal errors in the habitat class on which a point falls. Whilst some kind of buffering or filtering method could theoretically be used to obviate this obvious problem, it would be much better to have an expert decision and use a list of the principal habitats of a species.

10. The ability to be able to use more than three different clips at a time would be useful for future implementations of PDA.

11. One of the most advanced requirements for local priority areas analysis is the development of methods to order grid-cells according to biological and/or threat (and/or socio-economic) criteria and/or justifiable weightings thereof (for example, to order grid-cells according to

biological diversity divided by human population pressure). Criteria need to be expressed in terms of common currency for such orderings to be justifiable. Such methods could be implemented in a future PDA.

## Lessons learned

**Lesson 1:** The available data needs to be received from the participants, reformatted, integrated and cleaned in advance of a priority-setting workshop or the time available for analysis and interpretation and particularly input into the priority-setting process will be severely constrained.

**Lesson 2:** Large numbers of species cannot yet be processed by PDA (Arcview version) in real time by most computers, but one cannot count on automatic generation of shape files overnight, in case of errors. Large range polygon shape files must be generated well in advance on a powerful computer. For an internet server version of PDA this may not be such a limitation.

**Lesson 3:** Output of analytical results using the PDA query tools is currently not implemented and presently demands some work-around method. This must be borne in mind when considering preparation of analyses in advance

**Lesson 4:** Intelligent use of habitat files and other themes for generating polygon range files requires expert decisions beforehand on the appropriate habitats/attributes for species and is best NOT derived from the point data (except perhaps in the rare cases where the date and resolution of the habitat files correspond exactly to those for the points). However, PDA already offers the very useful facility of using habitat attributes in comma-delimited format provided they match the codes for the habitat file. This is the best approach at present, at least for groups where most of the important habitat types for each species are known.

## Acknowledgements

Thanks particularly to Lanto Andriamampianina (WCS) for help well beyond the call of duty during this workshop as well as to the many other helpful participants at PACT-Miray, ANGAP, WWF, CMP, etc.